

FEEDFORWARD PILOT-AIDED CARRIER SYNCHRONIZATION USING A DCT BASIS EXPANSION

Jabran Bhatti and Marc Moeneclaey

Department of Telecommunications and Information Processing
Ghent University
Belgium
{Jabran.Bhatti, mm}@telin.ugent.be

ABSTRACT

This contribution deals with phase noise estimation from pilot symbols. The phase noise process is approximated by an expansion of Discrete Cosine-Transform (DCT) basis functions containing only a few terms. We propose a feedforward algorithm that estimates the DCT coefficients without requiring detailed knowledge about the phase noise statistics. We demonstrate that the resulting (linearized) mean-square phase estimation error consists of two contributions: a contribution from the additive noise, that equals the Cramer-Rao lower bound, and a noise-independent contribution that results from the phase noise modeling error. We investigate the effect of the symbol sequence length and the number of estimated DCT coefficients on the estimation accuracy and on the corresponding bit error rate (BER). We propose a pilot symbol configuration allowing to estimate any number of DCT coefficients not exceeding the number of pilot symbols. For large block sizes, the DCT-based estimation algorithm substantially outperforms algorithms that estimate only the time-average or the linear trend of the carrier phase.

1. INTRODUCTION

Phase noise refers to random perturbations in the carrier phase, caused by imperfections in both transmitter and receiver oscillators. Compensation of this phase noise is critical since these disturbances can considerably degrade the system performance. The phase noise process typically has a low-pass spectrum [1]. Discrete-time processes that have a bandwidth which is considerably less than the sampling frequency can often be modeled as an expansion of suitable basis functions, that contains only a few terms. Such a basis expansion has been successfully applied in the context of channel estimation and equalization in wireless communications, where the coefficients of the channel impulse response are lowpass processes with a bandwidth that is limited by the Doppler frequency [2–4].

Several methods trying to tackle the phase noise problem exist:

- Phase noise can be tracked by means of a feedback algorithm that operates according to the principle of the phase-locked loop (PLL). Such algorithms give rise to rather long acquisition transients, as such they are not well suited to burst transmission systems [5, 6].
- The observation interval is divided into subintervals, and a feedforward algorithm is used to estimate within each subinterval the local time-average (or the linear trend) of the phase [6, 7]. However, in order that the piecewise constant (or linear) approximation of the phase noise be

accurate, the subintervals should be short, in which case a high sensitivity to additive noise occurs.

- Recently, iterative joint estimation and decoding/detection algorithms have been proposed that make use of the a priori statistics of the phase noise process [8, 9]. These algorithms are computationally rather complex, prevent the use of off-the-shelf decoders and assume detailed knowledge about the phase noise statistics at the receiver.

In this contribution, we apply the basis expansion model to the problem of phase noise estimation from pilot symbols only, using the orthogonal basis functions from the discrete cosine transform (DCT). In contrast to the case of channel estimation, the phase noise does not enter the observation model in a linear way. Section 2 presents the phase noise estimation algorithm, based on the estimation of only a few DCT coefficients. Section 3 contains the performance analysis of the proposed algorithm in terms of the mean-square error (MSE) of the phase estimate. Analysis results are confirmed by computer simulations in section 4, which consider both the MSE and the associated bit error rate (BER) degradation. Conclusions are drawn in section 5.

2. ESTIMATION ALGORITHM

We consider the transmission of a block of K data symbols over an AWGN channel that is affected by phase noise. The resulting received signal is represented as:

$$r(k) = a(k)e^{j\theta(k)} + w(k) \text{ for } k = 0, \dots, K-1 \quad (1)$$

where the index k refers to the k -th symbol interval of length T , $\{a(k)\}$ is a sequence of data symbols with symbol energy $E[|a(k)|^2] = E_s$, the additive noise $\{w(k)\}$ is a sequence of i.i.d. zero-mean circularly symmetric complex-valued Gaussian random variables with $E[|w(k)|^2] = N_0$, and $\theta(k)$ is a time-varying phase noise process with $K \times K$ correlation matrix \mathbf{R}_θ .

The symbol sequence $\{a(k)\}$ contains K_P known pilot symbols at positions k_i , $i = 0, \dots, K_P - 1$, with constant magnitude: $|a(k_i)|^2 = E_s$. From the observation of the received signal at the pilot symbol positions k_i , an estimate $\hat{\theta}(k)$ of the time-varying phase $\theta(k)$ is to be produced. The phase estimate will be used to rotate the received signal before data detection, i.e., the detection of the data symbols is based on $\{z(k)\} = \{r(k)\exp(-j\hat{\theta}(k))\}$. The detector is designed under the assumption of perfect carrier synchronization, i.e., $\hat{\theta}(k) = \theta(k)$. For uncoded transmission, the detection algorithm reduces to symbol-by-symbol detection:

$$\hat{a}(k) = \arg \min_{a \in A} |z(k) - a|^2, k \notin \{k_i, i = 0, \dots, K_P - 1\}$$

The authors wish to acknowledge the activity of the Network of Excellence in Wireless COMMunications NEWCOM++ of the European Commission (contract n. 216715) that motivated this work.

with \mathbf{A} denoting the symbol constellation. The phase $\theta(k)$ can be represented as a weighed sum of K basis functions over the interval $[0, K-1]$:

$$\theta(k) = \sum_{n=0}^{K-1} x_n \psi_n(k), \quad k = 0, \dots, K-1 \quad (2)$$

As $\theta(k)$ is essentially a lowpass process, it can be well approximated by the weighed sum of a *limited number* N ($\ll K$) of suitable basis functions:

$$\theta(k) \approx \sum_{n=0}^{N-1} x_n \psi_n(k), \quad k = 0, \dots, K-1 \quad (3)$$

In this contribution we make use of the orthonormal discrete cosine transform (DCT) basis functions, that are defined as

$$\psi_n(k) = \begin{cases} \sqrt{\frac{1}{K}} & n = 0 \\ \sqrt{\frac{2}{K}} \cos\left(\frac{\pi n}{K} \left(k + \frac{1}{2}\right)\right) & n > 0 \end{cases}$$

Hence, from (2), x_n is the n -th DCT coefficient of $\theta(k)$. As $\psi_n(k)$ has its energy concentrated near the frequencies $n/2KT$ and $-n/2KT$, the DCT basis functions are well suited to represent a lowpass process by means of a small number of basis functions.

In the following, we produce from the observation $\{r(k_i)\}$ at the pilot symbol positions k_i , with $i = 0, \dots, K_P - 1$, an estimate \hat{x}_n of the coefficients x_n , with $n = 0, \dots, N-1$, using the phase model (3) with equality. The final estimate $\hat{\theta}(k)$ is obtained by computing the inverse DCT of $\{\hat{x}_n\}$:

$$\hat{\theta}(k) = \sum_{n=0}^{N-1} \hat{x}_n \psi_n(k) \quad \text{for } k = 0, \dots, K-1 \quad (4)$$

However, as (3) is not an exact model of the true phase $\theta(k)$, the phase estimate is affected not only by the additive noise contained in the observation, but also by a phase noise modeling error. Considering the observations (1) at instants k_i , and assuming that (3) holds with equality, we obtain:

$$\mathbf{r}_P = \mathbf{D}(\mathbf{x})\mathbf{a}_P + \mathbf{w}_P \quad (5)$$

where, for $i = 0, \dots, K_P - 1$; $(\mathbf{r}_P)_i = r(k_i)$, $(\mathbf{w}_P)_i = w(k_i)$, $(\mathbf{a}_P)_i = a(k_i)$ and $\mathbf{D}(\mathbf{x})$ is a $K_P \times K_P$ diagonal matrix with

$$(\mathbf{D}(\mathbf{x}))_i = e^{j(\Psi_P \mathbf{x})_i}$$

and $(\Psi_P)_{i,n} = \psi_n(k_i)$, $(\mathbf{x})_n = x_n$, $n = 0, \dots, N-1$ with $N \leq K_P$. The $K_P \times 1$ vectors \mathbf{r}_P , \mathbf{a}_P and \mathbf{w}_P can be viewed as resulting from subsampling $\{r(k)\}$, $\{a(k)\}$ and $\{w(k)\}$ at the instants k_i that correspond to the pilot symbol positions. Similarly, the n -th column of the $K_P \times N$ matrix Ψ_P is obtained by subsampling the n -th DCT basis function $\psi_n(k)$. Maximum likelihood estimation of \mathbf{x} from \mathbf{r}_P results in

$$\hat{x}_{ML} = \arg \min_{\mathbf{x}} |\mathbf{r}_P - \mathbf{D}(\mathbf{x})\mathbf{a}_P|^2$$

As \mathbf{x} enters the observation \mathbf{r}_P in a non-linear way, the ML estimate is not easily obtained. Therefore, we resort to a suboptimum ad-hoc estimation of \mathbf{x} , which is based on the argument (angle) of the complex-valued observations. However, as the function $\arg(z)$ reduces the argument of z to an interval $[-\pi, \pi]$, taking $\arg(r(k_i))$ might give rise to phase

wrapping, especially when the time-average of $\theta(k)$ is close to $-\pi$ or π . In order to reduce the probability of phase wrapping, we first rotate the observation \mathbf{r} over an angle θ_{avg} that is close to the time-average of $\theta(k)$, then we estimate the DCT coefficients of the fluctuation $\theta(k) - \theta_{avg}$ and finally, we compute the phase estimate $\hat{\theta}(k)$. We select

$$\theta_{avg} = \arg \left(\sum_{i=0}^{K_P-1} r(k_i) \right)$$

and construct \mathbf{r}' with

$$\begin{aligned} (\mathbf{r}')_i &= r'(k_i) \\ &= \arg(r(k_i)a^*(k_i)\exp(-j\theta_{avg})) \\ \text{for } i &= 0, \dots, K_P - 1 \end{aligned} \quad (6)$$

We obtain an estimate $\hat{\mathbf{x}}'$ of the DCT coefficients of the fluctuation $\theta(k) - \theta_{avg}$ through a least-squares fit $\hat{\mathbf{x}}' = \arg \min_{\mathbf{x}} |\mathbf{r}' - \Psi_P \mathbf{x}|^2$, yielding:

$$\hat{\mathbf{x}}' = (\Psi_P^T \Psi_P)^{-1} \Psi_P^T \mathbf{r}' \quad (7)$$

The estimation of the phase trajectory involves the inversion of the $N \times N$ matrix $\Psi_P^T \Psi_P$, which depends on the pilot symbol positions $\{k_i, i = 0, \dots, K_P - 1\}$. In order that $(\Psi_P^T \Psi_P)^{-1}$ exists, we need $N \leq K_P$. Now we point out that the pilot symbol positions can be selected such that $\Psi_P^T \Psi_P$ is diagonal, or, equivalently, that the N columns of the $K_P \times N$ matrix Ψ_P are orthogonal. Such selection of $\{k_i\}$ avoids the need for matrix inversion in (7). Denoting by $\phi_n(i)$ the orthonormal DCT basis functions of length K_P , it is easily verified that selecting $\{k_i\}$ such that

$$k_i = \frac{iK}{K_P} + \frac{K - K_P}{2K_P}, \quad i = 0, \dots, K_P - 1 \quad (8)$$

gives rise to

$$\psi_n(k_i) = \sqrt{\frac{K_P}{K}} \phi_n(i) \quad \text{for } n = 0, \dots, K_P - 1$$

so that

$$\Psi_P^T \Psi_P = \frac{K_P}{K} \mathbf{I}_N \quad (9)$$

with \mathbf{I}_N denoting the $N \times N$ identity matrix. Equation (7) then reduces to

$$\hat{\mathbf{x}}' = \frac{K}{K_P} \Psi_P^T \mathbf{r}' \quad (10)$$

The phase estimate is given by

$$\hat{\theta} = \theta_{avg} \mathbf{1} + \frac{K}{K_P} \Psi_K \Psi_P^T \mathbf{r}' \quad (11)$$

where $(\hat{\theta})_k = \hat{\theta}(k)$, $(\mathbf{1})_k = 1$, $(\Psi_K)_{k,n} = \psi_n(k)$, $k = 0, \dots, K-1$; $n = 0, \dots, N-1$. Note from (11) that the estimation algorithm does not need specific knowledge about the phase noise process. As $r'(k_i)$ from (6) can be viewed as a noisy version of $\theta(k_i) - \theta_{avg}$, the phase estimate $\hat{\theta}$ from (11), or, equivalently, the phase estimate $\hat{\theta}(k)$ from (4), can be interpreted as an interpolated version of the subsampled noisy phase trajectory.

In order that all k_i from (8) be integer, K must be an

odd multiple of K_P , i.e. $K = (2d + 1)K_P$, yielding $k_i = (2d + 1)i + d$. The resulting pilot symbol configuration is suited for estimating any number of DCT coefficients not exceeding K_P . When K is not an odd multiple of K_P , rounding the right-hand side of (8) to the nearest integer gives rise to pilot symbol positions that still yield an essentially diagonal matrix $\Psi_P^T \Psi_P$ in which case the simplified equations (10) and (11) can still be used.

3. PERFORMANCE ANALYSIS

As the observation vector \mathbf{r}_P is a nonlinear function of the carrier phase, an exact analytical performance analysis is not feasible. Instead, we will resort to a linearization of the argument function in (6) in order to obtain tractable results. Linearization of the argument function yields

$$\begin{aligned} (\mathbf{r}'_i) &= \arg(e^{j(\theta(k_i) - \theta_{avg})}(E_s + a^*(k_i)w(k_i)e^{-j\theta(k_i)})) \\ &\approx \theta(k_i) - \theta_{avg} + n_P(i) \end{aligned} \quad (12)$$

for $i = 0, \dots, K_P - 1$, where $\{n_P(i)\}$ is a sequence of i.i.d. zero-mean Gaussian random variables with variance $N_0/2E_s$. Note that (12) incorporates the true phase $\theta(k_i)$ instead of the approximate model (3), so that our performance analysis will take the modeling error into account. In order that the linearization in (12) be valid, we need $|\theta(k_i) - \theta_{avg}| < \pi$ (because $|\arg(z)| < \pi$) and $|w(k_i)|^2 \ll E_s$; hence, the phase noise fluctuations should not cause phase wrapping and E_s/N_0 should be sufficiently large. Substituting (12) into (11) yields

$$\hat{\boldsymbol{\theta}} = \frac{K}{K_P} \Psi_K \Psi_P^T (\boldsymbol{\theta}_P + \mathbf{n}_P) \quad (13)$$

$$= \frac{K}{K_P} \Psi_K \Psi_P^T \mathbf{S} \boldsymbol{\theta} + \frac{K}{K_P} \Psi_K \Psi_P^T \mathbf{n}_P \quad (14)$$

where $(\mathbf{n}_P)_i = n_P(i)$, $(\boldsymbol{\theta}_P)_i = \theta(k_i)$ and the $K_P \times K$ matrix \mathbf{S} is such that its i -th row has a '1' at the k_i -th column and zeroes elsewhere ($i = 0, \dots, K_P - 1$). The estimation error resulting from (14) is given by

$$\hat{\boldsymbol{\theta}} - \boldsymbol{\theta} = \left(\frac{K}{K_P} \Psi_K \Psi_P^T \mathbf{S} - \mathbf{I}_K \right) \boldsymbol{\theta} + \frac{K}{K_P} \Psi_K \Psi_P^T \mathbf{n}_P \quad (15)$$

where \mathbf{I}_K denotes the $K \times K$ identity matrix. If the model (3) were exact, we would have $\boldsymbol{\theta} = \Psi_K \mathbf{x}$ and $\boldsymbol{\theta}_P = \Psi_P \mathbf{x}$, yielding

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta} + \frac{K}{K_P} \Psi_K \Psi_P^T \mathbf{n}_P$$

in which case the estimation error would be caused only by the additive noise.

As a performance measure of the estimation algorithm we consider the mean-square error (MSE), defined as

$$MSE = \frac{1}{K} E [\text{trace}((\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T)] \quad (16)$$

Substituting (15) into (16) yields

$$MSE = \frac{N_0}{2E_s} \frac{N}{K_P} + MSE_\infty \quad (17)$$

where

$$\begin{aligned} MSE_\infty &= \frac{K}{K_P^2} \text{tr} (E[\boldsymbol{\theta}_P \boldsymbol{\theta}_P^T] \Psi_P \Psi_P^T) \\ &\quad - \frac{2}{K_P} \text{tr} (E[\boldsymbol{\theta}_P \boldsymbol{\theta}^T] \Psi_K \Psi_P^T) \\ &\quad + E[\theta^2(k)] \end{aligned} \quad (18)$$

The first term in (17) denotes the contribution from the additive noise, whereas the second term in (17) constitutes a MSE floor, caused by the phase noise modeling error. The phase noise statistics affect the MSE floor through the autocorrelation matrix \mathbf{R}_θ . The MSE floor decreases with increasing N (because the modeling error is reduced when more DCT coefficients are taken into account), whereas the additive noise contribution to the MSE increases with N (because N parameters need to be estimated). Hence, there is an optimum value of N that minimizes the MSE.

From the nonlinear observation model (5), which assumes that (3) holds with equality, we compute the Cramer-Rao lower bound on the MSE (16) resulting from any unbiased estimate $\hat{\mathbf{x}}$ of the DCT coefficients of $\theta(k)$:

$$MSE \geq \frac{1}{K} \text{trace}(\mathbf{J}^{-1}) \quad (19)$$

In (19), \mathbf{J} denotes the Fisher information matrix related to the estimation of \mathbf{x} from (5), which is found to be:

$$(\mathbf{J})_{n,n'} = \frac{2E_s}{N_0} \frac{K}{K_P} \delta_{n-n'} \quad (20)$$

Combining (19) with (20) yields the following performance bound:

$$MSE \geq \frac{N_0}{2E_s} \frac{N}{K_P} \quad (21)$$

Expression (21) indicates that the sensitivity to additive noise increases with the number N of estimated DCT coefficients. Comparison of (17) and (21) shows that our ad hoc algorithm (11) yields the minimum possible (over all unbiased estimates) noise contribution to the MSE (assuming that the linearization of the observation model is valid).

4. SIMULATION RESULTS

In this section we assess the performance of the proposed technique in terms of the MSE of the phase estimate and the resulting BER degradation by means of computer simulations. We consider the presence of Wiener phase noise $\theta(k)$, which is described by the following system equation:

$$\theta(k+1) = \theta(k) + \Delta(k), k = 0, \dots, K-2 \quad (22)$$

where the initial phase noise value $\theta(0)$ is uniformly distributed in $[-\pi, \pi]$ and $\Delta(k)$ is a sequence of i.i.d. zero-mean Gaussian increments with variance σ_Δ^2 . Hence, $\theta(k)$ can be viewed as the output of an integrator with a white noise input. From (22) it follows that the variance of the Wiener phase noise increases linearly with the time index k , which indicates that the process is non-stationary.

First, we assume transmission of a block of length $K = 105$ symbols, consisting of $K_D = 90$ uncoded QPSK data symbols and $K_P = 15$ constant-energy pilot symbols ($\eta = K_P/K = 1/7$) that are inserted into the sequence according to (8):

- Figure 1 shows the MSE as a function of E_s/N_0 for $N = 1, 4$ and 10 , in the presence of Wiener phase noise with $\sigma_\Delta^2 = 0.0027 \text{ rad}^2$ (which corresponds to $\sigma_\Delta = 3^\circ$). We observe a MSE floor in the high- E_s/N_0 region which can be reduced by increasing the number N of estimated coefficients. Figure 1 also confirms that for low E_s/N_0 the MSE increases when N increases. This high- E_s/N_0 and low- E_s/N_0 behavior indicates that for given K , K_P and E_s/N_0 the MSE can be minimized by proper selection of N .

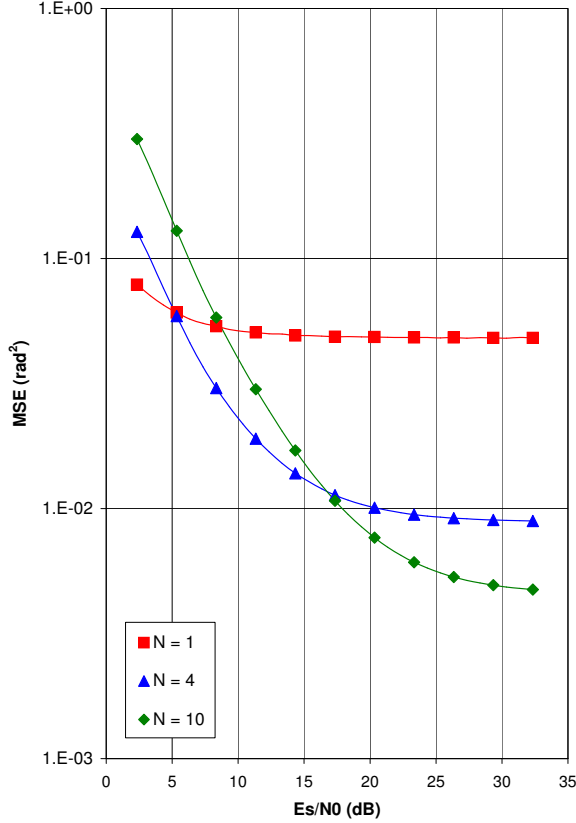


Figure 1: MSE when Wiener phase noise with $\sigma_{\Delta} = 3^{\circ}$ is present. $K = 105$, $K_P = 15$.

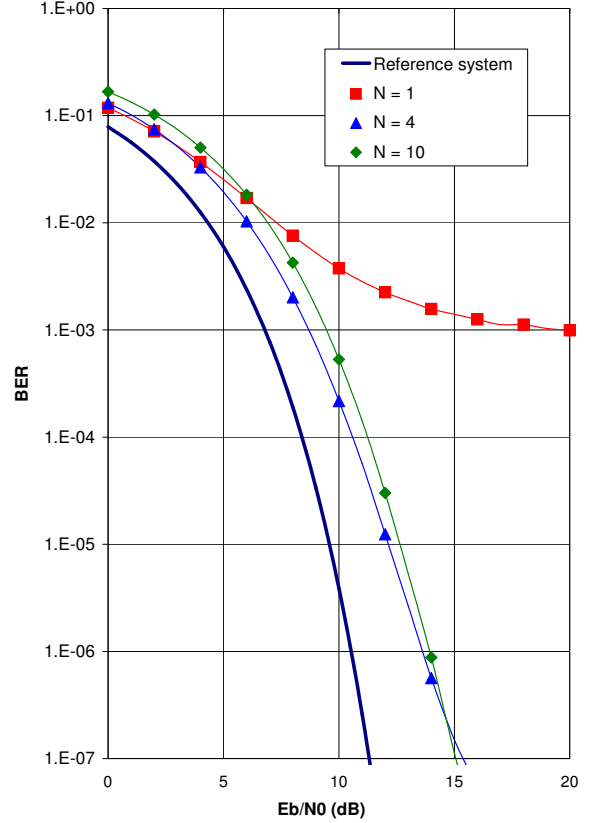


Figure 2: BER when Wiener phase noise with $\sigma_{\Delta} = 3^{\circ}$ is present. $K = 105$, $K_P = 15$.

- Figure 2 shows the Bit Error Rate (BER) as a function of E_b/N_0 for $N = 1, 4$ and 10 (E_b is the energy per transmitted bit, $E_s = 2(1 - \eta)E_b$ for QPSK). The reference BER curve corresponds to a system with perfect synchronization and no pilot symbols ($\eta = 0$). We observe that for low E_b/N_0 , it is sufficient to estimate only the time-average of the phase (i.e., $N = 1$). Estimating a higher number of DCT coefficients can lead to a worse BER performance for low E_b/N_0 , because the MSE of the phase estimate due to additive noise increases with N . At high E_b/N_0 a BER floor occurs which decreases with increasing N , so in this region it becomes beneficial to estimate more than just one DCT coefficient. Hence, the optimal number of estimated coefficients N_{opt} will depend on the operating E_b/N_0 .

Figure 3 shows the BER degradation¹ when (1) $\eta = 20\%$ and $\sigma_{\Delta} = 3^{\circ}$ and (2) $\eta = 10\%$ and $\sigma_{\Delta} = 2^{\circ}$, for the following phase noise estimation algorithms:

- The proposed DCT-based algorithm with orthogonal pilot symbol placement (8). The number of estimated coefficients N is chosen such that the BER degradation is minimum for the considered block length K .
- Estimation of only the time-average of the phase noise.

¹The BER degradation caused by some impairment is characterized by the increase (in dB) of E_b/N_0 (as compared to the case of no impairment) needed to maintain the BER at a specified reference level

Half the number of pilot symbols are arranged at the beginning of the symbol sequence, the remaining half are placed at the end of the sequence.

- The method from Luise et al. [7], where again half the number of pilot symbols are arranged at the beginning of the sequence and the other half at the end of the sequence. The phase noise over the total symbol block is approximated as a linear interpolation between the average phase values over the first and the second pilot symbol cluster.

We observe that estimating only the time-average or the linear trend of the phase noise yields poor BER performance, except for small K . For $K = 10$, the DCT-based algorithm also estimates the time-average only (because $N = 1$ is optimum for $K = 10$); we observe that the second algorithm (with pilot symbols at positions 0 and 9) performs slightly better than the DCT-based algorithm (with pilot symbols at positions 2 and 7) for $K = 10$. However, when the block length is increased, the DCT-based algorithm that estimates multiple DCT coefficients outperforms both the time-averaging algorithm and Luise et al.'s algorithm and leads to a BER degradation that decreases with increasing K until an optimal value for K is reached.

5. CONCLUSIONS AND REMARKS

In this contribution we have considered an ad hoc feedforward data-aided phase noise estimation algorithm that is based on the estimation of only a few (N) coefficients of the

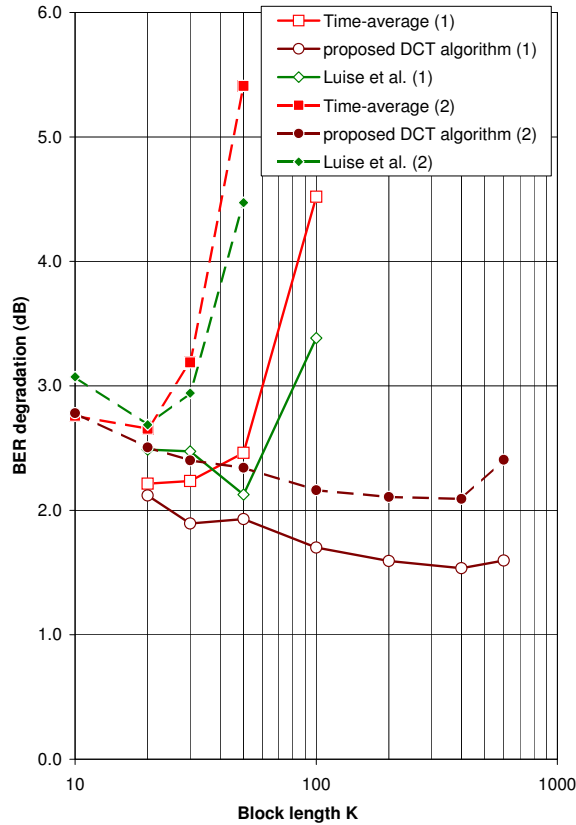


Figure 3: Comparison of BER degradation for $BER = 10^{-4}$ as function of K . (1) $\eta = 20\%$ and $\sigma_{\Delta} = 3^{\circ}$; (2) $\eta = 10\%$ and $\sigma_{\Delta} = 2^{\circ}$ for different estimation algorithms.

DCT basis expansion of the time-varying phase. The algorithm does not require detailed knowledge about the phase noise statistics. Linearization of the observation model has indicated that the mean-square error of the resulting estimate consists of an additive noise contribution (that increases with N) and a MSE floor caused by the phase noise modeling error (that decreases with N). The noise contribution coincides with the Cramer-Rao lower bound. These analytical findings have been confirmed by means of computer simulations. The numerical results illustrate that the MSE and BER degradation can be minimized by a suitable choice of K and N . For large K , substantial improvement is obtained as compared to the case where the phase noise is approximated by its time-average or a linear trend.

REFERENCES

- [1] A. Demir, A. Mehrotra, and Roychowdhury J., "Phase noise in oscillators: A unifying theory and numerical methods for characterization," *IEEE Transactions on Circuits and System - I: Fundamental Theory and Applications*, vol. 47, no. 5, pp. 655–674, May 2000.
- [2] G.B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," in *Proceedings of the IEEE*, 1998, vol. 86, pp. 1969–1986.
- [3] J.K. Tugnait and W. Luo, "Blind space-time multiuser channel estimation in time-varying DS-CDMA systems," *IEEE Transactions on Vehicular Technology*, vol. 55, pp. 207–218, January 2006.
- [4] O. Rousseaux, G. Leus, and M. Moonen, "Estimation and equalization of doubly selective channels using known symbol padding," *IEEE Transactions on Signal Processing*, vol. 54, pp. 979–990, March 2006.
- [5] U. Mengali and A.N. D'Andrea, *Synchronization Techniques for Digital Receivers*, Plenum Press, New York, USA, 1997.
- [6] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers : Synchronization, Channel Estimation and Signal Processing*, Wiley Series in Telecommunications and Signal Processing, USA, 1998.
- [7] L. Benvenuti, L. Giugno, V. Lottici, and M. Luise, "Code-aware carrier phase noise compensation on turbo-coded spectrally-efficient high-order modulations," in *8-th Intern. Work. on Signal Processing for Space Commun.*, Catania, Italy, September 2003, pp. 177–184.
- [8] G. Colavolpe, A. Barbieri, and G. Caire, "Algorithms for iterative decoding in the presence of strong phase noise," *IEEE Journal on selected areas in communications*, vol. 23, pp. 1748–1757, September 2005.
- [9] E. Panayirci, H. Cirpan, and M. Moeneclaey, "A sequential monte-carlo method for blind phase noise estimation and data detection," in *Proc. European Sign. Proc. Conf. (EUSIPCO)*, Antalya, Turkey, September 2005.